UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/660,635	09/13/2000	Alan Lawrence Higgins	D-ACD-W019 HIGGINS	7304

23122 7590 03/03/2005

RATNERPRESTIA
P O BOX 980
VALLEY FORGE, PA 19482-0980

EXAMINER	
WOZNIAK, JAMES S	

ART UNIT	PAPER NUMBER
2655	

DATE MAILED: 03/03/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/660,635	HIGGINS ET AL.	
	Examiner	Art Unit	
	James S. Wozniak	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *12/10/2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-9,11,12 and 15-18* is/are pending in the application.

4a) Of the above claim(s) *1-5* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *6-9,11,12 and 15-18* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *14 October 2004* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1. In response to the office action from 11/12/2004, the applicant has submitted a request for continued examination, filed 12/10/2004, amending the Claim 6, 7, 9, 11, 12, 15, 16, and 17, while adding Claim 18 and arguing to traverse the art rejection based on the limitations of Claim 6 and 18 *(Amendment, Pages 7-9)*. The applicant's arguments have been fully considered but are moot with respect to the new grounds of rejection in view of Häb-Umbach et al *(U.S. Patent: 5,873,061)* and Naito et al *(U.S. Patent: 5,983,178)*.

2. Based on the amendments to the drawings, specification, and claim 7, the examiner has withdrawn the previous objections directed towards minor informalities.

Claim Objections

3. **Claim 12** is objected to under 37 CFR 1.75(c), as being of improper dependent form for failing to further limit the subject matter of a previous claim. Applicant is required to cancel the claim(s), or amend the claim(s) to place the claim(s) in proper dependent form, or rewrite the claim(s) in independent form. In this case, the step of analyzing speech frames using a weighted Euclidean distance computation between the enrollment and test feature vectors is already recited in Claim 1.

4. **Claim 18** is objected to because of the following informalities: "Baum-Welsh" on Line 11 should be changed to --Baum Welch--.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. **Claims 6-9, 11-12, and 15-17** are rejected under 35 U.S.C. 103(a) as being unpatentable over Vysotsky et al (U.S. Patent: 5,832,063) in view of Gandhi et al (U.S. Patent: 5,687,287), in view of Kuhn et al (U.S. Patent: 6,343,267), and in yet further view of Häb-Umbach et al *(U.S. Patent: 5,873,061)*.

With respect to **Claim 6**, Vysotsky recites:

Causing said speaker to enroll by uttering from a vocabulary a predetermined number of combined words each word indicative of a number between one to nine *(user enrollment of a voice password comprising a string of digits, Col. 11, Lines 45-55)*.

Adapting the parameters of a set of word models for said vocabulary words based upon input speech data to provide adapted word models *(creating speaker dependent word models, Col. 8, Lines 40-43)*.

The step of analyzing the speech frames by comparison includes computing a weighted Euclidean distance between the enrollment feature vector and the test feature vector using a

discriminative analysis *(Euclidean distance used in speech recognition, Col. 8, Lines 44-47, and discriminative analysis of feature vectors, Col. 11, Lines 58-63).*

Vysotsky does not specifically suggest a bridging word "ti" and a feature vector as a concatenation of state mean vectors, however, Gandhi teaches the use of a bridging word and the concatenation of word models to form a feature vector in speaker verification *(use of the word "ty" in a user password, Col. 8, Lines 7-27, which is a functional equivalent of "ti," and the concatenation of word models to form a feature vector to be used for speaker enrollment and verification, Col. 4, Line 53- Col. 5, Line 35).*

Vysotsky and Gandhi are directed towards a similar field of endeavor in speaker recognition, and would have been obvious for combination in order to obtain the capability of recognizing an entire spoken password sequence for speaker verification *(Gandhi, Col. 3, Lines 37-40, and Col. 6, Lines 6-18).*

Vysotsky in view of Gandhi do not teach the ability to convert variable input to fixed-length feature vectors that are independent of the order of words spoken or the speaking rate, however Kuhn discloses:

Ability to convert variable input to fixed-length feature vectors *(dimensionality reduction, Col. 6, Lines 62-64, and Col. 7, Lines 23-26)* that are independent of the speaking rate *(adaptive speaker models in the form of a supervector that is fully populated with parameter values for recognizing speech, Col. 9, Lines 41-51. Also, using a maximum likelihood technique, a supervector is selected which is most consistent with input speech, Col. 9, Lines 4-13, so that, regardless of the rate of speech, a proper supervector would be selected for speaker and speech*

recognition. Furthermore, using singular value decomposition the supervector dimensionality is reduced, Col. 6, Lines 36-45).

Vysotsky, Gandhi, and Kuhn are analogous art because they are from a similar field of endeavor in speech and speaker recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to combine the use of a supervector containing all speech parameter values within an Eigenspace for speaker adaptation as taught by Kuhn with the speaker verification method through feature extraction and the computation of Euclidean distances between feature vectors as taught by Vysotsky to provide for adaptive speaker recognition if speech rate is altered, especially in the enrollment process since all speakers do not speak at the same rate of speed, or if the order of words has been altered, for example, if the order of numbers within a voice password is changed .

Vysotsky, Gandhi, and Kuhn do not teach that the test feature vector is based on extracting the same predetermined number of combined words in the test speech data as in the enrollment data, however, Häb-Umbach teaches sub-word units extracted from input speech that can be used for speech recognition in an alternate order *(subsequent utterance for recognition which, within the scope of the disclosed invention, could utilize a sequence of registered sub-word units from a training utterance in a different order, Col. 5, Lines 51-56)*;

Vysotsky, Gandhi, Kuhn, and Häb-Umbach are analogous art because they are from a similar field of endeavor in speech recognition, and thus are obvious in combination for the benefit of providing a more memory efficient recognition dictionary for speaker verification *(Häb-Umbach, Col. 1, Lines 36-45)*.

With respect to **Claim 7**, Vysotsky teaches speaker verification through feature extraction and computing Euclidean distances between feature vectors as applied to Claim 6, which also utilizes a method of HMM adaptation through a Gaussian estimation *(Col. 8, Lines 40-43)*. Vysotsky does not specifically further suggest that the aforementioned method of Gaussian estimation utilizing a Baum-Welch algorithm, however it would have been obvious to one of ordinary skill in the art, at the time of invention, to specifically utilize a Baum-Welch algorithm for HMM parameter adaptation since it is a well-known and common means of HMM parameter estimation in the art of speech recognition and has readily available software.

With respect to **Claim 8**, Gandhi further recites:

The predetermined number of vocabulary words comprises five words, namely, "four", "six", "seven", "nine", and "ti" *(voice password for user verification as a string of any number of digits comprising a password, Col. 8, Lines 7-27)*.

With respect to **Claim 9**, Gandhi further recites:

The enrollment and test feature vectors are created by concatenating the state-mean vectors of the adapted HMM word models (*the concatenation of word models to form a feature vector to be used for speaker enrollment and verification, Col. 4, Line 53- Col. 5, Line 35)*.

With respect to **Claim 11**, Vysotsky further discloses:

Comparing said enrollment feature vector obtained from said enrollment with the test feature vector obtained from a speech test to determine the identity of a test speaker voice *(voice verification through comparison of feature vectors corresponding to a voice password to identify either a true or impostor speaker, Col. 11, Lines 45-63)*.

With respect to **Claim 12**, Vysotsky further recites:

Feature comparison is implemented by subjecting said vectors to a weighted Euclidean Distance computation *(Euclidean distance used in speech recognition, Col. 8, Lines 44-47)*.

With respect to **Claim 15**, Vysotsky further discloses a feature vector matrix used for comparison to input speech feature vectors for voice identification *(Col. 8, Lines 4-7)*. Also, it would have been obvious to one of ordinary skill in the art, at the time of invention that the dimensionality of this matrix could have a value of 1568, for instance, in a 49X32 or other such matrix configuration, based on desired system settings.

With respect to **Claim 16**, Vysotsky further discloses:

Forming said feature vector for each speaker using the difference in vectors between a first and second speaker channel *(speaker reference model adapted and thus formed according to changes in speaker and channel coupling, Col. 11, Line 64- Col. 12, Line 1)*.

With respect to **Claim 17**, Vysotsky teaches the speaker verification system featuring speaker enrollment through a voice password and adaptive word models responsive to changes between speaker channels as applied to Claim 16. Vysotsky does not specifically suggest the approximation of speech with white noise channel differences in deriving speaker features as recited in Claim 17, however it would have been obvious to one of ordinary skill in the art, at the time of invention, to include white noise approximation in speech features, since white noise is common to telephone communication channels, and thus, should be included within the speaker feature vectors modeled from speech inputs to the communication channels to better approximate their expected characteristics.

7. **Claim 18** is rejected under 35 U.S.C. 103(a) as being unpatentable over Häb-Umbach et al *(U.S. Patent: 5,873,061)* in view of Vysotsky et al (U.S. Patent: 5,832,063), and further in view of Naito et al *(U.S. Patent: 5,983,178)*.

With respect to **Claim 18**, Häb-Umbach discloses:

Forming enrollment speech data as a first plurality of pair-phrases using a set of words, the set of words consisting of a predetermined number of words *(sub units of an utterance obtained from an speech input to and stored for future speech recognition, Col. 2, Line 67- Col. 2, Line 1, and Col. 3, Lines 18-34)*;

Forming test speech data as a second plurality of pair phrases from the same set of words, the second plurality of pair-phrases different from the first plurality of pair-phrases *(subsequent utterance for recognition which, within the scope of the disclosed invention, could utilize a sequence of registered sub-word units from a training utterance in a different order, Col. 5, Lines 51-56)*;

Ordering the first set of adapted HMM word models into a first sequence *(sequence of sub-word models, Col. 5, Lines 46-49)*;

Ordering the second set of adapted HMM word models into a second sequence, the second sequence and the first sequence having the same order and the same predetermined number of words *(matching a subsequent utterance to a sequence of previously registered sub-word unit models for recognition, Col. 5, Lines 51-56)*; and

Häb-Umbach does not specifically suggest that whole words can be used in place of sub-word units, however Vysotsky teaches the interchangeability of whole words and sub-word units *(Col. 7, Lines 58-62)*, wherein whole word units can be considered to be digits within a password

string *(speaker verification utilizing a password comprising a 7-10 digit string, Col. 11, Lines 45-55)*.

Häb-Umbach and Vysotsky are analogous art because they are from a similar field of endeavor in word recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Häb-Umbach with the interchangeability of whole words and sub-word units, wherein whole word units are considered to be digits within a password string as taught by Vysotsky in order to implement the speech recognition method that can recognize a word even if that particular utterance has not been previously registered as taught by Häb-Umbach in a practical application for verifying speaker identity for access to a system via a voice password *(Vysotsky, Col. 11, Lines 56-63)*, thus increasing method usability and providing a more memory efficient recognition dictionary for speaker verification *(Häb-Umbach, Col. 1, Lines 36-45)*.

Although Häb-Umbach teaches the comparison between an enrollment and test utterance for recognition as noted above, Häb-Umbach does not teach that the comparison is based on Euclidean distance nor the use of a Baum-Welch algorithm to form adapted HMM models prior to model ordering, however both of these processes are well known in the speech recognition art, as is evidenced by Naito *(Col. 9, Line 62- Col. 10, Line 28)*.

Häb-Umbach, Vysotsky, and Naito are analogous art because they are from a similar field of endeavor in speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Häb-Umbach in view of Vysotsky with the use of Euclidean distance comparison and the Baum-Welch algorithm for

speaker identification as taught by Naito provide adapted HMMs that can provide a higher speech recognition rate *(Naito, Col. 1, Lines 60-65)*.

Conclusion

8. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure:

Matsui et al *(U.S. Patent: 5,835,890)*- teaches the transformation of a speech input into a sequence of feature parameters utilized in the recognition of a four digit number sequence.

9. Any inquiry concerning this communication or earlier communications from the examiner should be directed to James S. Wozniak whose telephone number is (703) 305-8669 and email is James.Wozniak@uspto.gov. The examiner can normally be reached on Mondays-Fridays, 8:30-4:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached at (703) 305-4827. The fax/phone number for the Technology Center 2600 where this application is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the technology center receptionist whose telephone number is (703) 306-0377.

James S. Wozniak
2/17/2005

DAVID L. OMETZ
PRIMARY EXAMINER